Supplement to

The AMERICAN PHYSICS TEACHER

VOLUME 3 | SEPTEMBER, 1935 | NUMBER 3, SUPPLEMENT

The report of the Committee on Tests of the American Association of Physics Teachers again is issued as a Supplement in order that it may reach readers early in the new academic year. As before, funds to help defray the expenses of publication have been supplied by the Cooperative Test Service and the Committee on Educational Testing of the American Council on Education.

Perhaps the most important and interesting feature of the present report is the comparison of pre-study and post-study scores. This program still remains the largest college testing project in progress and through it the Association, and physicists as a group, can without doubt make a contribution of major importance both to physics teaching and to higher education as a whole. It is hoped that departments will continue their cooperation during the present year.—THE EDITOR.

FEB 11 1936 A. H. COMPTON

The 1934–1935 College Physics Testing Program

I. INTRODUCTION

THE present report marks the second year in which the American Association of Physics Teachers has sponsored a nation-wide program of testing in elementary college physics. Although the report is shorter than that of last year, its brevity is justified on the grounds that it can be read most intelligently as a supplement to the earlier report and as an extension of it. Technical matters concerning test construction and the analysis of the tests as measuring devices have been omitted to conserve space; and the specific uses to which the participating departments have put test results have been waived in favor of a subsequent report devoted entirely to this purpose.

The program has obviously merited the wholehearted support of the Association, despite minor criticisms. That only 269 departments ordered tests as against 355 for last year indicates no loss of interest. The program for 1935 was more elaborate and called for pre-study and post-study testing which many colleges were unable to carry out for lack of time. Correspondence offering suggestions and helpful criticisms continues unabated. Hence, the Committee will continue the program for another year and will elaborate it still more. This extension will include pre- and post-study tests in all six of the topics so far used, mechanics, heat, sound, light, electricity, and modern physics. It is recommended that departments use as many of these separately printed tests, both before and after course work, as time and content of courses will permit. Pre-study tests will be labeled Form A; post-study tests, Form B. Orders should be directed to the Cooperative Test Service, 437 West 59th Street, New York City.

The personnel of the Committee for the coming year will remain unchanged except for the addition of Dr. Harvey Lemon of the University of Chicago, whose long experience with physics tests will make him a valuable aid.

Participating Colleges

Alabama
Woman's College of Alabama*

Arizona
Arizona State Teachers College
University of Arizona*

Arkansas
A. and M. College*
College of the Ozarks
Harding College*
University of Arkansas

California
Bakersfield Junior College*
Citrus Junior College
College of the Pacific*
Lassen High School*
Occidental College*
Pacific Union College*
Pasadena Junior College
Riverside Junior College*
San Bernardino Valley Junior College*
University of Redlands

Canada
St. Francis Xavier College*
University of British Columbia

Colorado
Colorado College*
Grand Junction State Junior College*
State Teachers College, Greeley*

Connecticut
Connecticut State College
U. S. Coast Guard Academy
Wesleyan University

District of Columbia
Catholic University*
Mt. Vernon Seminary*
Wilson Teachers College*

Florida
Southern College
University of Florida

Georgia
Agnes Scott College*
Emory University

Idaho
University of Idaho

Illinois
Armour Institute of Technology
Aurora College
Central Y. M. C. A. College
DePaul University*
Eureka College
Greenville College
Illinois College
Joliet Junior College
Knox College*
Lake Forest College
LaSalle-Peru-Oglesby Junior College
Lyons Junior College*
Millikin University*
Morton Junior College
Northern Illinois State Teachers College*
Springfield Junior College*
University of Chicago
University of Illinois
Western Illinois State Teachers College

Indiana
Butler University*
DePauw University
Earlham College*
Goshen College*
Manchester College
Purdue University*
University of Notre Dame*

Iowa
Buena Vista College
Central College*
Cornell College
Fort Dodge Junior College
Graceland College
Grinnell College*
Iowa State College
Parsons College*
Simpson College*
St. Ambrose College*
University of Dubuque*
Waldorf College*
Wartburg College
Western Union College*

Kansas
Baker University*
Central College
College of Emporia
Friends University
Garden City Junior College*
Hutchinson Junior College
Kansas State College*
Kansas Wesleyan University*
McPherson College*
St. Benedict's College*

Kentucky
Asbury College
Berea College
Eastern Kentucky State Teachers College
Union College*
University of Louisville
Villa Madonna College*

Louisiana
Louisiana College*
Tulane University*

Maine
Bates College
Bowdoin College*
Colby College
University of Maine
Westbrook Junior College*

Maryland
Loyola College
Western Maryland College*
Woodstock College

Massachusetts
Clark University*
Harvard University*
Mount Holyoke College*
Simmons College
Wellesley College

Michigan
Alma College
Battle Creek College*
Bay City Junior College*
Central State Teachers College*
Highland Park Junior College*
Jackson Junior College*
Kalamazoo College
Michigan State Normal College*
Port Huron Junior College
Western State Teachers College

Minnesota
Bemidji State Teachers College*
College of St. Catherine*
College of St. Thomas
Concordia College*
Duluth Junior College*
Duluth State Teachers College*
Ely Junior College
Eveleth Junior College
Hibbing Junior College
Itasca Junior College
Macalester College
Rochester Junior College*
St. John's University*
St. Mary's College
State Teachers College, St. Cloud*

Mississippi
Mississippi State College*

Missouri
Culver Stockton College*
Kemper Military School*
Lindenwood College
Moberly Junior College
Park College
University of Missouri
Washington University
Westminster College*
William Jewell College

Montana
Montana State Normal College*
Northern Montana College

Nebraska
Doane College
Municipal University of Omaha
Nebraska Wesleyan University
Union College*
University of Nebraska*
York College*

New Jersey
New Jersey State Teachers College
St. Peter's College
Upsala College*

New Mexico
New Mexico Military Institute
New Mexico State College*

New York
Alfred University*
Bard College*
Colgate University
Columbia College
Cornell University*
Hamilton College
Houghton College*
Hunter College
Long Island University
Niagara University
Packer Collegiate Institute
St. Lawrence University*
Sarah Lawrence College*
Skidmore College
Syracuse University*
University of Rochester
Vassar College*
Wells College*

North Carolina
Catawba College
Duke University*
Guilford College*
Lees-McRae College*
Queens-Chicora College*
Western Carolina Teachers College*
Woman's College, University of North Carolina

North Dakota
Jamestown College
North Dakota School of Forestry*
State Teachers College, Minot
State University of North Dakota

Ohio
Baldwin Wallace College
Bowling Green State College
Defiance College
Heidelburg College*
John Carroll University*
Miami University*
Urbana Junior College
Xavier University*

Oklahoma
Murray State School of Agriculture*
Panhandle A. and M. College*
University of Oklahoma
University Preparatory School*

Oregon
Linfield College
Reed College*
University of Oregon*

Pennsylvania
Allegheny College
Bryn Mawr College
College Misericordia*
Duquesne University*
Elizabethtown College*
Geneva College
Haverford College
Immaculata College*
Lafayette College
LaSalle College*
Muhlenberg College*
Pennsylvania State College
St. Joseph's College
St. Thomas College*
St. Vincent College
Swarthmore College
Thiel College*
University of Pennsylvania
Ursinus College*
Waynesburg College
West Chester State Teachers College*
Wilson College

Rhode Island
Providence College

South Carolina
Benedict College*
College of Charleston*
Presbyterian College

South Dakota
Augustana College
Huron College
Northern State Teachers College*

Tennessee
A. and I. State College*
David Lipscomb College*
Maryville College
Southwestern
Tennessee Wesleyan College
Union University
University of the South*
University of Tennessee

Texas
A. and M. College of Texas*
Amarillo Junior College*
Brownsville Junior College*
Daniel Baker College*
John Tarleton Agricultural College*
Paris Junior College
Sam Houston State Teachers College*
Schreiner Institute
Texas Christian University
Texas Lutheran College*
Texas Technological College
University of Houston*
University of Texas*
West Texas State Teachers College*

Utah
Brigham Young University
University of Utah

Vermont
Green Mountain Junior College*

Virginia
Hampton Institute
Mary Baldwin College*
University of Virginia

Washington
College of Puget Sound
Grays Harbor Junior College*
State College of Washington*
University of Washington
Walla Walla College

West Virginia
Concord State Teachers College
Storer College

Wisconsin
Carroll College*
Lawrence College*
Milton College
St. Norbert College*
State Teachers College, La-Crosse
State Teachers College, Milwaukee*
State Teachers College, Oshkosh*
State Teachers College, River Falls
State Teachers College, Superior

Participating colleges

The number of colleges that ordered tests for the 1935 physics program is 269. Of these, 137 did not send in their results in time for inclusion in this report, and some colleges did not use all the topical tests. Thus, the total number of colleges represented in the succeeding tables and charts is 132. The complete list of colleges that ordered tests is given by states on the preceding page. Those not represented in any of the tables or charts are starred.

The tests

The tests used in the 1935 program closely parallel, both in form and content, those used in the 1934 program. Since the latter were fully described in the report of the 1934 program, it is unnecessary to give details here.[1] The report of the 1935 test results which follows will be given in approximately the same form and the same language as the 1934 report.

II. The Test Results

National percentiles for men and women separately

In the report of the 1934 program, national percentiles were given for all students who took the tests without regard to sex. In answer to

[1] Am. Phys. Teacher 3, Supplement, 132–136 (1934).

Table I. *National percentiles for men.**

	M	H	S	M+H	M+H+S	L	E	MP	L+E	L+E+MP	
No. Cases	4962	4887	3631	4458	3074	3619	3903	1510	3340	1424	
Mean	18.8	14.4	7.5	33.4	40.8	12.2	17.9	4.6	30.0	35.4	
Sigma	9.6	5.8	3.4	14.4	17.2	5.8	7.6	3.4	12.3	14.3	
Percentile											Percentile
100	51	30	16	78	94	35	43	19	73	88	100
99	44	27	15	69	83	27	37	14	63	74	99
98	41	26	(15)	66	79	25	35	13	58	69	98
97	39	25	14	63	76	24	33	12	55	66	97
96	38	(25)	(14)	61	74	23	32	11	54	64	96
95	36	24	13	60	72	22	(32)	(11)	52	62	95
94	35	(24)	(13)	58	70	(22)	31	(11)	51	60	94
93	34	23	(13)	57	69	21	30	10	50	59	93
92	(34)	(23)	12	55	66	(21)	(30)	(10)	49	58	92
91	33	(23)	(12)	54	(66)	20	29	(10)	48	57	91
90	32	22	(12)	53	65	(20)	28	9	47	55	90
88	30	21	(12)	51	63	19	27	(9)	45	53	88
86	29	(21)	11	49	61	(19)	(27)	8	44	51	86
84	28	20	(11)	48	58	18	26	(8)	43	50	84
82	(28)	(20)	(11)	47	57	(18)	25	(8)	41	48	82
80	27	(20)	10	45	55	17	(25)	(8)	40	47	80
75	25	19	(10)	43	52	16	23	7	38	44	75
70	23	18	9	41	49	15	22	6	36	41	70
65	22	17	(9)	38	47	14	20	(6)	34	39	65
60	20	16	8	36	44	13	19	5	32	37	60
55	19	15	(8)	34	42	12	18	(5)	30	36	55
50	18	14	7	33	40	(12)	17	4	29	34	50
45	17	(14)	(7)	31	37	11	16	(4)	27	32	45
40	16	13	6	29	35	10	15	3	26	30	40
35	14	12	(6)	27	33	(10)	14	(3)	24	29	35
30	13	11	5	25	30	9	13	2	23	27	30
25	12	10	(5)	23	28	8	12	(2)	21	25	25
20	10	9	4	21	25	7	11	(2)	19	23	20
18	(10)	(9)	(4)	20	24	(7)	(11)	1	(19)	22	18
16	9	8	(4)	19	23	6	10	(1)	18	21	16
14	(9)	(8)	(4)	18	22	(6)	(10)	(1)	17	20	14
12	8	7	3	16	21	(6)	9	(1)	16	19	12
10	7	(7)	(3)	15	19	5	(9)	0	15	18	10
9	(7)	6	(3)	14	18	(5)	8	(0)	(15)	(18)	9
8	6	(6)	(3)	13	(18)	4	(8)	(0)	14	17	8
7	(6)	(6)	2	(13)	17	(4)	(8)	(0)	(14)	16	7
6	5	5	(2)	12	16	(4)	7	(0)	13	(16)	6
5	4	(5)	(2)	11	15	3	(7)	(0)	12	15	5
4	(4)	4	1	10	14	(3)	6	(0)	11	14	4
3	3	(4)	(1)	9	13	2	5	(0)	10	13	3
2	2	3	(1)	8	11	(2)	4	(0)	9	12	2
1	0	2	0	5	9	0	3	(0)	7	10	1

* The scales are based upon returns for students tested after studying the various topics, and show true percentiles, calculated from the distributions of scores available at the time of computation. Each score in each column shows the upper score limit of the percentile indicated at the extreme right and left of the line. For example, the bottom entry in the column for total score on M+H shows that all scores of 5 or below have a percentile value of 1; all scores of 6, 7, and 8 have a percentile value of 2; and all scores above 69 have a percentile value of 100. Since colleges used varying combinations of tests, the numbers of cases vary from column to column. The mean and sigma of the scores are shown at the top of each column. When a score appears on a scale more than once, use the figure not in parenthesis.

TABLE II. *National percentiles for women.**

	M	H	S	M+H	M+H+S	L	E	MP	L+E	L+E+MP	
No. Cases	711	693	458	662	393	482	446	136	304	128	
Mean	14.7	11.6	7.3	26.4	33.9	12.3	14.9	3.7	25.4	28.0	
Sigma	7.5	5.1	3.1	11.6	13.7	6.0	6.1	2.8	10.4	11.4	
Percentile											Percentile
100	53	30	15	83	84	33	39	12	57	63	100
98	31	23	13	55	69	26	30	11	50	52	98
93	26	20	12	44	56	21	24	8	41	46	93
90	24	18	11	41	52	20	23	(8)	39	43	90
84	22	17	(11)	38	47	18	21	7	36	40	84
80	20	16	10	35	45	17	20	6	34	38	80
75	19	15	(10)	33	41	16	19	5	32	36	75
70	18	14	9	31	40	15	17	(5)	30	34	70
60	16	13	8	28	37	13	16	4	27	30	60
50	14	12	7	25	33	12	14	3	24	27	50
40	12	10	(7)	23	30	10	13	(3)	22	23	40
30	11	9	6	20	27	9	12	2	20	22	30
25	9	8	5	19	25	8	11	(2)	18	21	25
20	8	7	(5)	17	22	7	10	1	17	20	20
16	(8)	6	4	15	20	6	9	(1)	15	18	16
10	6	5	3	12	16	5	8	0	13	14	10
7	5	4	2	11	14	4	7	(0)	11	11	7
3	2	3	1	8	11	2	4	(0)	9	7	3
1	0	1	(1)	4	8	0	3	(0)	5	3	1

* Because of the small numbers of cases, the percentiles are shown in abbreviated form. The table is to be read in the same manner as Table I.

numerous suggestions, and in view of the sex differences found in the 1934 program, the Committee presents separate percentile tables for men and for women students. In view of the smaller numbers of women students, and the unknown selective factors that may have operated to produce this sampling, the differences between Table I, for men, and Table II, for women, should be interpreted with caution. For the same reasons, the percentiles for women students, Table II, are given in abbreviated form.

Tables I and II enable each college to find the percentile rank on each topic or combination of topics for each student among all the students of either sex who took the tests. From these tables, it is evident that the tests are well adjusted in difficulty for the participating colleges. The scores are not piled up at either end of the scale; hence, there is room at both the top and bottom for differentiating between good and poor students on all topical tests, with the possible exception of that for MP. Even in this test, however, only about 10 percent of either sex receive zero scores, which is probably due to the fact that in some colleges this topic receives little or no emphasis in the elementary physics course.

Variability of achievement

The fifteen colleges in Fig. 1 were chosen to represent the whole range of medians from highest to lowest and all types of institutions reporting results in time for inclusion in the report. The fifteen institutions include engineering schools, universities, four-year colleges, junior colleges, and teachers' colleges. Fig. 1 shows two types of variability; the first is the variability of median scores in individual colleges. The differences between colleges are very large indeed. So far as the functions measured by the combined M, H, and S tests are concerned, the lowest group of colleges has almost nothing in common with the three or four colleges at the high end of the scale; however, the variability within colleges suggests that adjustment of course and pace to the needs and abilities of individual students must be made within the college, despite mitigation of the problem by progressive pre-college selection and guidance of students.

Fig. 2 illustrates for combined scores on L, E, and MP what Fig. 1 shows for M, H, and S. Both are to be interpreted in the same way. Figs. 1 and 2 do not include the same colleges. It is obvious from both of these figures that the students in some colleges learn much less physics of the type measured by these tests than those in other colleges. In college 15, Fig. 1, for example, all of the students are below the national average, while in college 1 over 90 percent are above; and in Fig. 2 approximately similar relations are found.

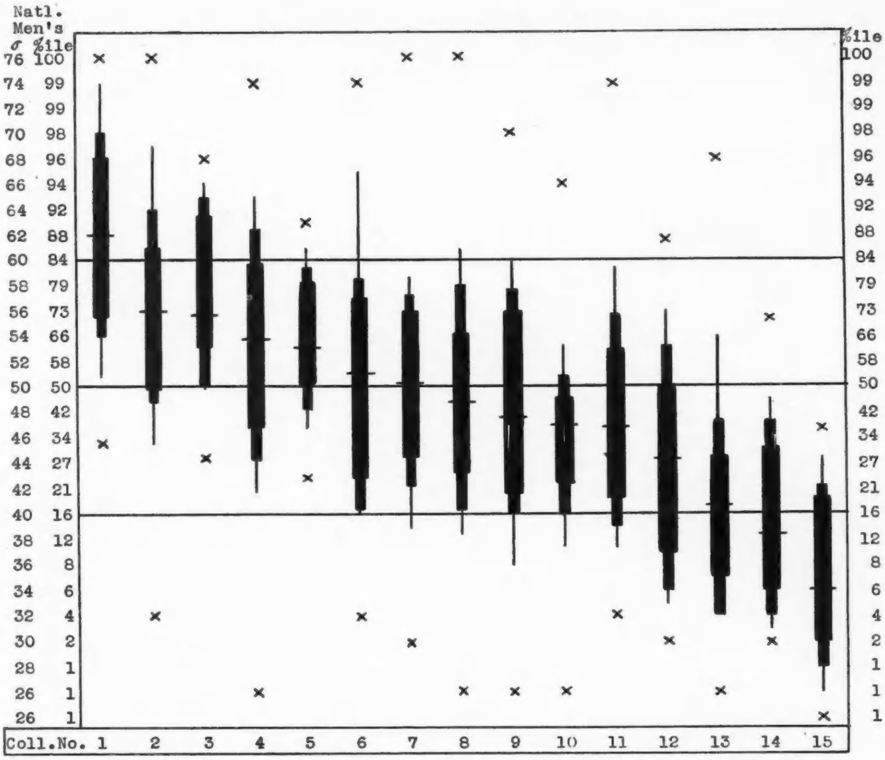


FIG. 1. Variability of achievement as measured by the combined scores on M, H, and S. The middle horizontal line shows the national median and the other two are at the 16th and 84th percentiles of the national distribution. Each of the bars represents an individual college. The wide portion of each bar represents the range of scores of the middle half in each college. The narrow parts extend to the 16th and 84th percentile in each college, i.e., one standard deviation above and one below the mean. The lines at the ends extend down to the 10th percentile and up to the 90th percentile. The crosses below the bars represent the lowest scores and those above represent the highest scores in the several colleges (the range). The short cross line at the middle of each bar represents the median score of the college. Although this chart is based entirely on percentiles, the scale has been altered roughly to a sigma scale, so that vertical distances are approximately comparable. The sigma scale is derived from the percentile scale. The bars represent both men and women students, but the percentile scale is that for men students from Table I.

In Fig. 1, half the students in college 15 are in the lowest 6 percent of the national distributions for men; in college 1 half the students are in the highest 12 percent. Part of these differences may be due to curricular differences; but even if these differences were based on completely valid tests for all colleges, it would still not mean that the low-average college is necessarily inferior to the high in its *educational* contribution. Even though the achievement of its students *in physics* may be clearly inferior, the college may be giving its students other experiences more valuable to them, in view of their abilities, interests, and social needs, than mastery of physics.

College averages

The fact that variability of scores within each college group is much more striking than differences between groups makes it necessary to interpret any distribution of college averages with

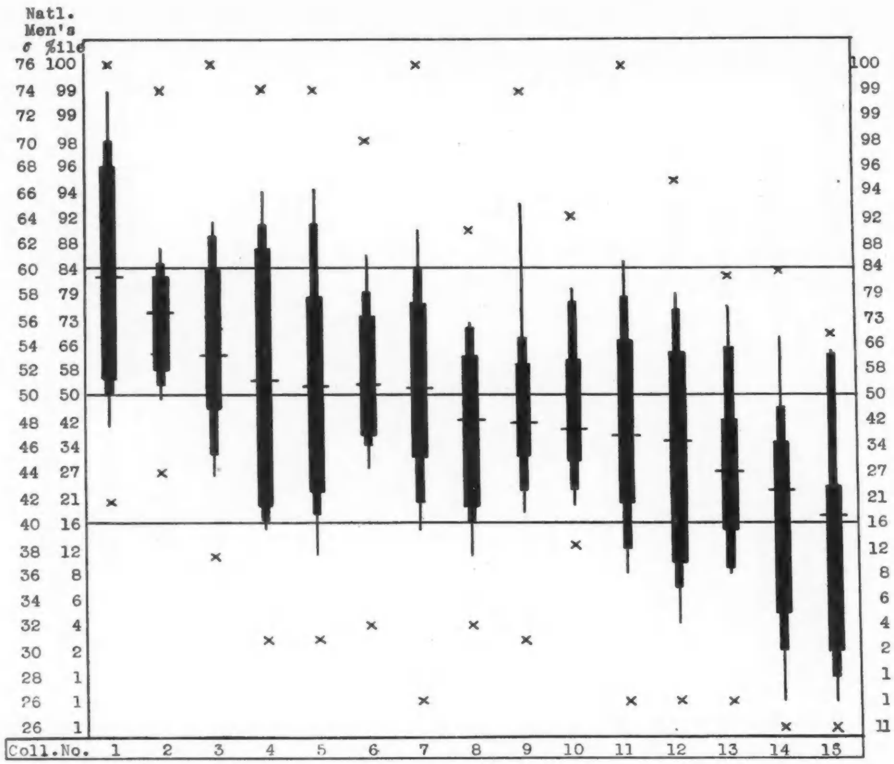


FIG. 2. Variability of achievement as measured by combined scores on L, E, and MP. To be read in the same manner as Fig. 1.

caution. Recent reports[2] from the North Central Association make it clear that college accreditation in the future will likely be along more constructive lines than in the past. College objectives and selection of students vary greatly. The North Central Association holds that these facts should be taken into account for accreditation. The same principle, we believe, should operate among physics departments. A physics department should not necessarily regard itself as superior to any other merely on the basis of a higher average score on the physics tests. The whole trend of the present study suggests that the best grounds for the relative ranking of departments might well be in terms of how well they diagnose the needs of students, how well they provide opportunity for the progress of their students at self-determined, differential rates, and how adequately they define and attempt to meet their avowed objectives. Hours and units and credits and average scores furnish an excellent basis for self-study but are not remotely adequate for ranking the departments in order of merit for accreditation and like purposes. It is quite possible that some colleges of low average scores may be meeting the educational and cultural needs of their particular students more effectively than some colleges of high average score. The work of a college cannot be fairly judged except by taking into account the abilities and total educational needs of its own student groups.

Nevertheless, with this word of caution,

[2] North Central Association of Colleges and Secondary Schools, Commission on Institutions of Higher Education, *Statement of Policy* (Feb. 3, 1934). M. E. Haggerty, "Accrediting Institutions of Higher Education," North Central Assoc. Quar. **9**, 177 (1934). G. F. Zook, "Accrediting Schools and Colleges," Ed. Rec. **15**, 10 (1934):

TABLE III. *Distribution of college averages.**

Group	*Mechanics*							*Heat*							*Sound*						
	1	2	3	4	5	6	Total	1	2	3	4	5	6	Total	1	2	3	4	5	6	Total
National Percentiles																					
88–91	1		1		1		**3**														
84–87			1			1	**2**											1			**1**
79–83			1			1	**2**			2				**2**	1		1		1		**3**
72–78			2		1		**3**	1		3	1	1		**6**			2			1	**3**
66–71			4	2		1	**7**			1	1	1	2	**5**	1		2	2	2	1	**8**
58–65	1		5	2	2	3	**13**	1		11			5	**17**	1	1	8	1		1	**12**
50–57	4	1	10		3	2	**20**	5	1	7		3	1	**17**	2	1	6	1	1	2	**13**
42–49	2		9	1		1	**13**	1	1	8	2	1		**13**	4	2	9	1		3	**19**
34–41		4	6	1	3		**14**	3	2	7	2	3	1	**18**	1		4	1	1		**7**
27–33	2		4	2			**8**			6	3	1	1	**11**		1	7	2	2		**12**
21–26	1		3	1		1	**6**	1	3	4			1	**9**	1		5	1	1		**8**
16–20	2	1	1	1		1	**6**	1			2			**3**	2	1	1				**4**
12–15	1	1					**2**	1						**1**			1			1	**2**
8–11			1	1		1	**3**			1				**1**		1					**1**
6–7																	1				**1**
	14	7	48	11	10	12	**102**	14	7	50	11	10	11	**103**	13	7	47	10	8	9	**94**

Mechanics+Heat							*Mech.+Heat+Sound*							Group
1	2	3	4	5	6	Total	1	2	3	4	5	6	Total	
														National Percentiles
		1				**1**			1				**1**	88–91
1		1		1		**3**	1		1			1	**3**	84–87
		1			1	**2**			1				**1**	79–83
		5				**5**			4	1	1	1	**7**	72–78
		1	2	2	·2	**7**			4		1	1	**6**	66–71
1		5		1	2	**9**	1		4	1	2	2	**10**	58–65
3		8	1	3	1	**16**	3		8			1	**12**	50–57
2	2	10	1			**15**	2	2	5		1		**10**	42–49
1	2	6	1	3	1	**14**	2	2	7	2	1		**14**	34–41
1	2	4	3			**10**		1	4	2	1	1	**9**	27–33
1		1			1	**3**	1		2				**3**	21–26
1		1	1		1	**4**	1			2		1	**4**	16–20
1	1		1			**3**		1					**1**	12–15
						1								8–11
									1				**1**	6–7
12	7	45	10	10	9	**93**	11	6	42	8	7	8	**82**	

Group	*Light*							*Electricity*							*Modern Physics*						
	1	2	3	4	5	6	Total	1	2	3	4	5	6	Total	1	2	3	4	5	6	Total
National Percentiles																					
88–91			1		1		**2**														
84–87			1				**1**														
79–83	1		2				**3**			2		1		**3**	2	1	1				**4**
72–78		2				1	**3**	1		3			3	**7**				1			**1**
66–71	2			1		1	**4**	1		5			1	**7**			2	1		1	**4**
58–65	4		7	2	2	1	**16**	3	1	5	3	1		**13**			8			1	**9**
50–57			9	3	3	4	**19**	4		5		3	3	**15**	2		3	1	1	2	**9**
42–49	2	1	7				**10**	3	1	10	3		2	**19**	2	1	2		1		**6**
34–41	1	1	7			1	**10**		3	8	2		1	**14**		2	3			3	**8**
27–33	3	1	5	1		1	**11**	1		3	1		1	**6**	2		1				**3**
21–26	1	1	5	1		1	**9**	2	1	5	1	1		**10**			3			1	**4**
16–20			2	1	1		**4**			1			1	**2**							
12–15		1	1				**2**		1	1				**2**							
8–11				1			**1**	1	1	2				**4**							
6–7			1				**1**			1				**1**							
	14	7	48	10	7	10	**96**	16	8	51	10	6	12	**103**	8	4	23	3	2	8	**48**

Light+Electricity							*L+E+MP.*							Group
1	2	3	4	5	6	Total	1	2	3	4	5	6	Total	
														National Percentiles
				1		**1**								88–91
		2				**2**	1						**1**	84–87
1		1			1	**3**			1				**1**	79–83
		2			1	**3**	1		1			2	**4**	72–78
2	1	4	1			**8**	1		1			1	**3**	66–71
3	1	4	2	3	2	**15**			1	2	1		**4**	58–65
2		9	2	1	2	**16**	1		4		1	2	**8**	50–57
1	2	4			1	**8**	1		6			1	**8**	42–49
1		11			2	**14**	1	2	3				**6**	34–41
3		3	3			**9**			3	1		1	**5**	27–33
	1	2				**3**								21–26
		3	1	1	1	**6**			1			1	**2**	16–20
	2					**2**		1	1				**2**	12–15
		1				**1**								8–11
		1				**1**								6–7
13	7	47	9	6	10	**92**	6	3	22	3	2	8	**44**	

* The means on all six subjects and on four combinations are here distributed in terms of national percentiles. The first six distributions in each topic section are for types of institutions as follows:

1. Men's Liberal Arts Colleges
2. Women's Liberal Arts Colleges
3. Coeducational Liberal Arts Colleges
4. Teachers Colleges
5. Agricultural and Engineering Colleges
6. Junior Colleges

The last column in each section (in bold face) gives the distribution of averages for all types of institution combined. Although this distribution is based on percentiles, the scale has been altered so that the intervals correspond approximately to a sigma scale. The vertical distances are therefore roughly comparable.

distributions of college averages on the physics tests are presented in Table III. Any college, by referring its average scores to this table, may identify its relative position among the participating groups.

Table III differs from the corresponding table in the 1934 report in that the distributions of averages are given for types of colleges as well as for all colleges together. Differences between types should be interpreted cautiously, since the sampling of certain types is obviously inadequate. So far as the samplings go, however, it is apparent that the averages of all types on most of the topical tests vary almost as widely as individual pupil scores vary in any one college represented in Figs. 1 and 2.

Gains in achievement

The test results of those colleges administering tests before and after study[3] were analyzed to discover the average gain of all students, without

[3] Post-study percentile tables are given in Tables I and II. Pre-study tables are not given in this report, but may be found in the Preliminary Report on the 1935 College Physics Testing Program, which was sent to all participating colleges.

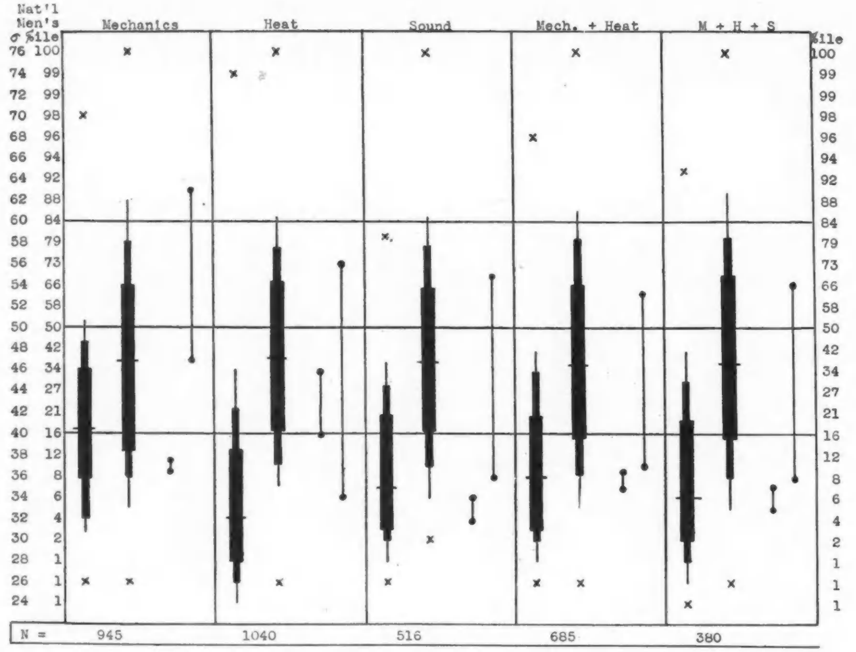


FIG. 3. Pre- and post-study comparisons for M, H, and S. The two bars at the left of each section represent distributions of scores of all students who took the tests both before and after study. The two following lines connect dots representing the pre- and post-study means for the two colleges having the smallest and largest gains in each topic.

regard to type of school, and of student groups in colleges making the largest and smallest gains.

The amount of gain within single colleges varies markedly. These differences, however, are not subject to direct comparison, since the abilities, aptitudes, and needs vary from college to college. The average gain of all students is of general interest; the gain of any one is of interest only to that college. It must be kept in mind that the average class gain must be interpreted in the light of local conditions. Each college must set up its own norms of expectancy and the deviations from these norms may be analyzed with reference to the individual case. Individual differences cover a wider range and are educationally more significant than the variations occurring among groups. Charting of single case gains cannot be done in this study; it may be done, however, by each department to determine differences in growth. Doubtless some such charts will be startling. Within the same class, some students will make tremendous gains in achievement, others will make little or none. Those deviations which are most marked may be investigated with a view to further clinical diagnosis and educational guidance. But the average gain for the school is most significant when it is taken as the point of reference for the interpretation of the individual gain.

In this connection, it is difficult to exaggerate the importance of the variability of scores in the pre-study groups. Teachers in physics and other subjects are only too painfully aware of the variability in post-study groups because there are too few classes in which some students are not "flunked"; but not many teachers are clearly aware of the highly important and encouraging fact that a few students at the *beginning* of a course are already able to equal or exceed the average of the whole class at the *end* of the course. Fig. 3 shows that in M and in H, at least one student at the beginning was in

the highest five percent of the class at the end of the semester; and in nearly all topics 10 percent of the students *at the beginning* secured scores which were above the average of the class *at the end*. In M about 30 percent of the pupils secured pre-study scores above the post-study average score of the entire group that took the test.

The advantages of early identification of such promising pupils are obvious and it is gratifying to note that during the last few years the search for and special study of such pupils has been a growing concern of an increasing number of teachers. Of course, no single test can be expected to give a fully trustworthy index of even one of the many qualities required for a productive career in physics, but students who *before* formal instruction secure higher scores than the average student secures *after* formal instruction are at least worthy of being studied. Some of them richly reward the teacher who vouchsafes them special attention and encouragement.

The students at the other end of the scale also merit special attention. Some of them after taking a full course, at the expense of the state or of their parents, are still in the score-range of the lowest 1 percent of pre-study norms. Some of these students work harder than the majority of those who make large gains and passing grades. Although it is obvious that they do not deserve, or rather that their lot will not be improved (and might be aggravated) by passing grades *in physics*, it is the opinion of a growing number of teachers that such students deserve something more constructive and helpful for their time and money than a mere failing grade. It is equally true that society deserves more for the time and money spent on such pupils than the stigmatizing of them as "failures." The question of what to do for such students can be answered only by careful and long continued experimentation, in which physicists will play their part. An essential part of such experimentation will be a surer identification of such students and a more adequate study of their growth along various lines of development. Growth cannot be studied without comparable measures; hence the emphasis which the committee has placed on the importance of developing *comparable* tests in physics.

III. Relations of Four Selective Factors to Test Scores

The four factors here studied are (1) the prerequisites for admission to the elementary college physics course, (2) the study of physics in high school, (3) the number of hours of credit allotted to the college physics course, and (4) the professional goal of the students as reported by them on the test cover-page. The relation of each of these to test scores is presented in terms of the average test scores of students in each indicated category. Since the relation of each factor is studied independently, without reference to other factors that may be associated with it in systematic or chance fashion, the indications must be viewed as merely suggestive. In the matter of prerequisites, for example, the indications may be largely or entirely a function of the sampling of colleges involved, rather than of the influence of the particular group of prerequisites. Nevertheless, since such factors as prerequisites and hours of credit normally operate categorically on entire college physics class enrollments, the relations or absence of relations presented below may have considerable significance in the reconsideration of college policies regarding these and other similar factors.

Prerequisites[4]

Quantitative data will not be given extensively on the results of studying prerequisites. The differences obtained were too largely unreliable or at best insigificant. Verbal statements about the findings seem all that can be justified. The indications are but little different from those obtained last year.

The average scores on each variate in group 1 were nearly identical. It seems almost inescapable that other factors not dependent upon the specific sets of stated prerequisites have more influence on the averages than the prerequisites

[4] Prerequisites were grouped as follows (numbers in the text refer to these groups):

1. None; Algebra and plane geometry; Algebra, trigonometry, and physics
2. Algebra, plane geometry, and trigonometry; Algebra, plane geometry, and physics; Algebra and trigonometry
3. Algebra
4. Algebra, geometry, trigonometry, and physics
5. Algebra and physics
6. Algebra, geometry, trigonometry, calculus, and physics
7. Trigonometry
8. Trigonometry and physics
9. Algebra, geometry, trigonometry, and calculus

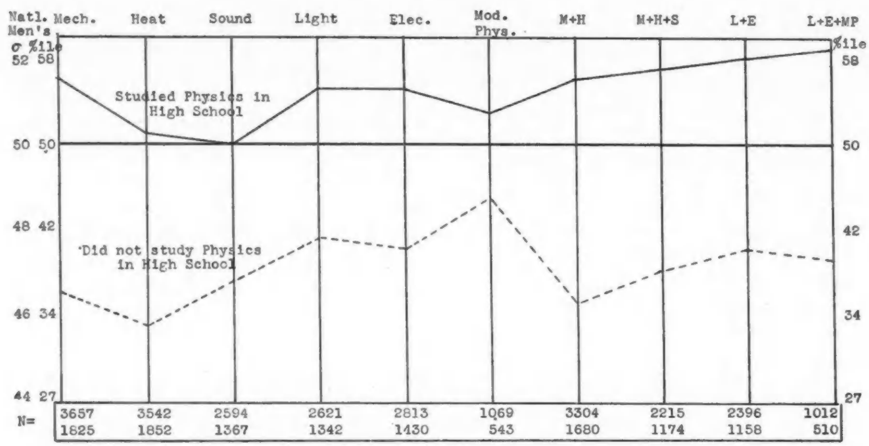


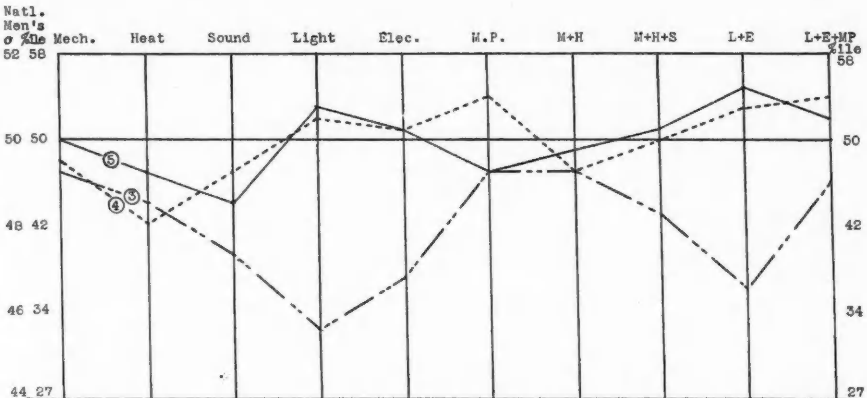
FIG. 4. Averages of students who studied and did not study physics in high school. The data are graphed in terms of the national percentile scale for men, altered, as in Fig. 1, to conform approximately to the sigma scale. Vertical distances are, therefore, roughly comparable.

themselves. In group 2, the uncertainties are increased by the small and varying numbers of cases. But if we compare groups 1 and 2, involving from 709 to 1952 cases, their relations do indicate that there is a consistent difference in favor of the second group of prerequisites. Group 6, representing the prerequisite of high school physics and mathematics through calculus, stands appropriately at the top, but group 9, involving the same prerequisites except high school physics, stands at the bottom in three of the topics and near the top in two topics.

While, as already indicated, the relations among the several groups are by no means conclusive, the Committee cannot escape the suggestion that if the prerequisites were really enforced, the differences in favor of the higher prerequisites would be both greater and more consistent. The common current method of enforcing these and all other prerequisites is to require that entrants to the physics course shall have "credits" in the prerequisite subjects. Twenty years of research have shown that "credits" do not, and under present conditions cannot, have any definite meaning. Figs. 1 and 2 above show that even in physics college credits do not have a definite meaning: approximately as large a proportion of the students in the colleges at the right of Figs. 1 and 2 received "college credit" in elementary physics as of the students in the colleges in the upper left corner of Figs. 1 and 2. The practical suggestion which the Committee offers at this point is that more colleges should inaugurate the systematic use of placement tests with all freshmen and other new students entering college, and follow these up with comparable achievement tests in all college subjects in which such comparable tests are available. Many colleges are already using placement and achievement tests with increasingly good results, decreasing wasted time and energy for both teachers and students.

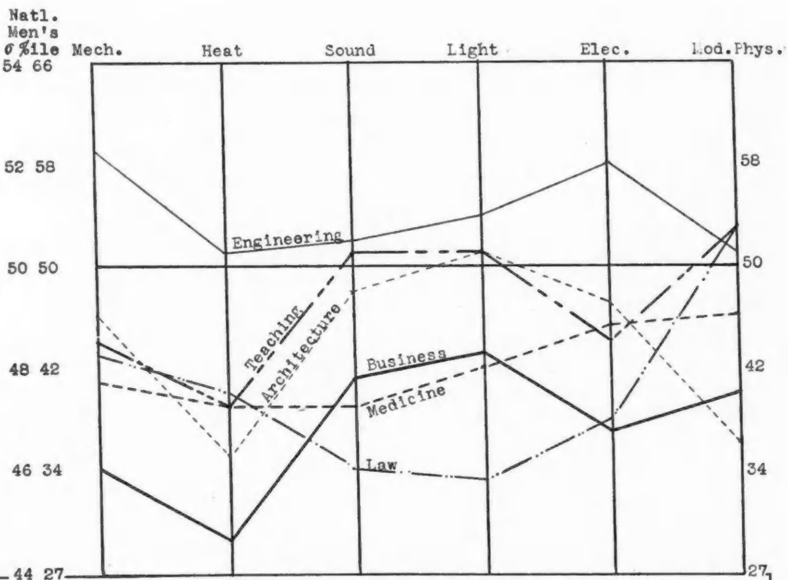
Physics in high school

Fig. 4 presents a violent contrast to the confusion in the prerequisite groupings described above. Here we have two fairly large groups of students. The group that studied physics in high school is consistently at or above the national average, while the other group is consistently below. The average difference is nearly one-half a standard deviation. If this figure may be taken at face value, it is apparent that the study of physics in high school tends toward higher achievement in college physics; but whether it is the longer experience with physics, or the selective influence of college physics following high school physics, or both, cannot be resolved by these data.



N	Mech.	Heat	Sound	Light	Elec.	M.P.	M+H	M+H+S	L+E	L+E+MP
3	1119	1094	474	380	395	145	1046	455	353	118
4	2003	2072	1435	1400	1604	560	1839	1253	1160	533
5	2278	2195	1934	2097	2070	883	2025	1596	1931	842

FIG. 5. Averages of students who received 3, 4, and 5 hours of college credit in physics, graphed in terms of the national percentile scale for men. The scale is the same as in Fig. 4.



No. of Cases:	Mech.	Heat	Sound	Light	Elec.	Mod.Phys.
Engineering	1825	1758	1107	1134	1298	565
Teaching	675	676	523	505	467	147
Architecture	70	68	49	39	42	20
Business	252	243	186	182	193	50
Medicine	1285	1275	1049	999	1034	361
Law	79	87	63	48	55	20

FIG. 6. Averages of professional goal groups. The scale is the same as in Fig. 4.

TABLE IV. *Difficulty and validity indices for each item in the M, H, and S, 1935B test forms, and L, E, and MP, 1935 test forms.**

Item No.	*Mechanics*		*Heat*		*Sound*		*Light*		*Electricity*		*Modern Physics*	
	Diff.	Validity	Diff.	Validity	Diff.	Validity	Diff.	Validity	Diff.	Validity	Diff.	Validity
1	53	5	51	2	81	6	85	7	94	5	55	5
2	72	5	81	4	27	5	85	3	94	6	44	5
3	59	5	57	6	69	4	77	4	88	5	42	3
4	59	4	93	6	70	5	74	6	74	7	47	3
5	68	7	73	6	87	7	69	2	83	5	61	4
6	62	5	45	5	67	6	58	4	88	1	66	4
7	71	8	56	6	32	3	51	4	62	5	61	7
8	65	6	86	5	26	1	87	4	63	5	12	1
9	30	4	67	4	78	2	77	5	79	0	38	4
10	49	6	56	4	64	3	51	5	69	5	38	7
11	44	3	60	9	54	3	57	5	61	4	43	4
12	80	4	66	6	46	5	39	4	60	6	47	5
13	58	6	47	3	40	4	35	3	64	4	18	1
14	41	6	77	4	38	5	75	4	49	4	27	3
15	56	7	31	6	44	5	50	0	60	5	37	3
16	61	4	67	5	36	4	50	6	62	5	13	4
17	47	6	75	4			69	5	46	7	34	3
18	67	3	41	6			48	4	40	3	10	1
19	42	6	58	7			54	1	49	8	14	2
20	71	4	32	4			34	3	60	7		
21	44	4	28	3			78	5	56	6		
22	44	1	69	9			25	4	43	4		
23	73	4	20	3			43	4	50	6		
24	16	4	79	2			25	6	52	5		
25	29	6	41	4			31	4	47	6		
26	45	7	58	2			45	5	55	10		
27	42	9	71	5			18	4	48	3		
28	44	4	19	4			47	5	59	7		
29	42	6	23	7			29	5	40	5		
30	56	5	41	6			27	3	45	5		
31	31	2					47	5	27	2		
32	55	9					6	0	36	6		
33	63	6					18	1	38	8		
34	25	1					32	7	50	6		
35	28	5					18	4	51	4		
36	45	7							28	3		
37	32	7							23	5		
38	13	8							20	5		
39	35	5							23	5		
40	26	4							26	3		
41	23	4							12	0		
42	38	7							29	6		
43	28	4							20	3		
44	21	5							22	2		
45	42	2										
46	31	5										
47	36	3										
48	75	5										
49	43	4										
50	27	7										
51	73	3										
52	42	4										
53	16	4										
54	5	3										

* The test papers analyzed were those of students who had taken all three tests. The difficulty ratings are percentages of students who answered each item correctly. The validity index indicates the degree to which an item discriminates between the high- and low-scoring students. The groups of high and low students were selected on the basis of total scores for the tests of each semester. A validity index of **0** means that as many poor students answered the item correctly as good students. Indices of 2 indicate satisfactory items; indices of 4 or above are very good items.

Hours of college credit

Fig. 5 indicates that the 3-hour group is inferior in Sound, Light, and Electricity, but otherwise there is little to choose among the 3-, 4-, and 5-hour credit groups. These indications are clearly worthy of further and more analytical study, especially when we remember that the averages graphed here conceal enormous variability within each of the groups. Fig. 3 showed that some students at the beginning of the course are already better than the average student at the end. To give such students the same "credit," even though in most cases it is with a grade of A instead of C, merely perpetuates the travesty of making college credits "rubber yardsticks" of ability and achievement. The Committee again urgently recommends the systematic use of comparable achievement tests as a partial basis for assigning credits.

Professional goal groups

Fig. 6 indicates, quite within expectations, that the group looking forward to engineering is higher in physics than any other pre-professional group. The pre-business group is lowest in all topics save S and L. The pre-law group, small

and variable in numbers, is generally low, except in MP. The pre-medical group is low in most topics, but approaches the national average for men in E and MP. The lines for the architecture and teaching groups run the same irregular course across the chart, but part company with a large difference in MP.

Here again the indications are irregular, and clouded by several uncertainties, of which the most important are sampling, local differences in specific emphases in the physics course, and the unreliability of the basis of grouping. In most if not all of these groups it is quite probable that large proportions of the students will never enter the indicated professions, either because they are unable to finish the professional course, or because they shift to some other profession more appropriate to their interests. But the differences, after discounting all vitiating factors, are large enough to be suggestive of guidance possibilities along the line of establishing more reliable norms for various professional groups, and of adapting the content of science courses more specifically to the actual needs of the different professional groups.

IV. Structure of the Tests: Item Analysis

Difficulty and validity of test items

Table IV shows the difficulty and validity indices for each item or question in each topical test (Form 1935B in M, H, S; Form 1935 in L, E, MP). The difficulty index is in terms of the percentage of students that answered the question correctly. Thus Item 1 in M was answered correctly by 53 percent of the students; Item 1 in H by 51 percent; Item 1 in S by 81 percent; etc.

The validity index is in terms of the difference between the proportions of good and poor students who answered each item correctly, this difference being expressed in quarter-sigma units as explained in the report of the 1934 program.[1] Thus Mechanics Item 1 is a very good item; the difference between the proportions of good and poor students that answer it correctly is five quarter-sigmas. In other words, it differentiates the good from the poor students very notably. It is suggested that physics teachers study these indices by transcribing them to the appropriate items on a copy of the tests.

Table V. *Summary from Table IV of difficulty and validity indices.*

Difficulty	M	H	S	L	E	MP
90		1			2	
85		1	1	3	2	
80	1	1	1		1	
75	1	3	1	4	1	
70	5	2	1	1	1	
65	3	4	2	2	1	1
60	3	1	1		8	2
55	6	5		2	3	1
50	1	1	1	5	4	
45	4	2	1	4	6	2
40	11	3	2	1	3	3
35	3		2	2	2	3
30	4	2	1	3		1
25	6	2	2	4	4	1
20	2	1			5	
15	2	1		3		1
10	1				1	4
5	1			1		
Validity	**M**	**H**	**S**	**L**	**E**	**MP**
10					1	
9	2	2				
8	2				2	
7	7	2	1	2	4	2
6	9	8	2	3	8	
5	10	4	5	9	14	3
4	15	8	3	12	5	5
3	5	3	3	4	5	5
2	2	3	1	1	2	1
1	2		1	2	1	3
0				2	2	

Summary of difficulty and validity indices

Table V is provided to facilitate an appraisal of the topical tests as wholes from the viewpoint of the difficulty and validity of the items. The upper part of the table shows that the items in five topics are very well distributed in difficulty. In MP no question was answered correctly by as many as 70 percent, and four of the 19 items were answered correctly by about 10 percent of the students. These data tend to confirm the judgment of several physics teachers that the items of this topical test are too hard. The Committee will try to secure a better distribution in next year's MP test.

The lower part of Table V shows a satisfactory or high degree of validity in nearly all the items in all topics. None of the questions is negative; that is, none is answered correctly more frequently by poor than by good groups of students. Only four of the 198 questions have zero indices; that is, only four are answered correctly by equal proportions of poor and good groups. Only 13 of the 198 questions have indices lower than +2, which means that 185 of the 198 questions show differences as great as one-half sigma or more in favor of the good student group. About 150 of the 198 questions have indices of +4 or greater; that is, favor the good students by one

sigma or more. These indications are very gratifying, and encourage the Committee to make further efforts along the same lines that have been followed thus far.

V. Conclusions

The Committee again extends thanks to the departments that have aided in conducting the program. Of particular significance are the reports on uses of test results that have been volunteered in answer to the suggestion made last year. As has already been said, a summary of these reports has not been given herewith; but later in the present year a special publication is planned, to be devoted entirely to this matter. Departments that have not as yet made contributions are invited, again, to do so before December 1, 1935, if possible. Letters or studies on uses of test scores should be sent to *F. S. Beers, Strahan House, University of Georgia, Athens, Georgia.*

Although the present report has several features which are elaborations or refinements of the earlier study, scientifically supported generalizations cannot be drawn from the data here presented. As in 1934, readers are urged to regard the concluding section of the report as suggestive rather than as conclusive.

One of the most striking inferences to be drawn from the data is this: a high average score on the tests is no guarantee that a department is doing its best for students enrolled in physics courses. Before any judgments can be made with reference to the efficiency of departments, it is necessary that the objectives of teaching in the light of student ability and needs be defined. In turn, the objectives of teaching cannot well be determined apart from a thorough-going diagnosis of student ability and interest. For this reason it is recommended that data from physics tests be interpreted in the light of placement and general ability examinations. Most colleges use such tests to a greater or less extent at the opening of the fall term.

Since the testing program in physics is the most elaborate measurement project in any given subject matter field, physicists have an unusual opportunity, if not a challenge, to influence local administrations. It is the opinion of many impartial observers that knowledge about students —their achievement, their interests—is a crying need. Physicists can do much to bring this point of view home to college executives and thus contribute to better selection and classification of students and to the shaping of curricula in keeping with the powers and interests of all classes of matriculants.

In this connection, the Committee suggests that every physicist interested in the policy of his own college read "Examinations Old and New: Their Uses and Abuses," by Max McConn, in the *Educational Record*, October, 1935.

Among other inferences which can be drawn from the data of the report is this: the examinations used in the program measure, and not in completely satisfactory fashion, only one of the desired outcomes of physics courses—namely, intellectual achievement. These examinations should not be used as the sole criterion of achievement but should be supplemented by tests made and administered locally. In addition, tests and methods of inquiry should be devised for the purpose of evaluating personal characteristics of students. It is as important that a prospective employer know something about the personalities of men or women whom he may wish to hire as that he know how well they understand their professional work. For example, a young graduate with a major in physics may be excellent in research projects; but because he dislikes being in close association with other people and does not speak with ease, he may be wholly unfitted to take a job as a teacher.

Other indications of the report suggest numerous additional interpretations along these lines. Only one more, however, will be mentioned. When the preliminary report on the program appeared early in the spring of 1935, it included a tabulation which emphasized the predictive value of the tests. This tabulation has been omitted from the present report except as its indications may be inferred from Fig. 3. The preliminary report carried a column designating gains in terms of gross scores for various levels of achievement in the pre- and post-study scales. From the figures given it is apparent that students doing well on pre-study tests do better on post-study tests than those making a poor initial showing. True, the initially poor student with a

ranking of 10 may have improved to a ranking of 20, which is a 100 percent gain. Conversely, the student making a rank of 80 and a gain of 20 has gained but 25 percent. It is obvious, however, that the first student knows very little about physics whereas the second one probably knows a good deal. Gains expressed as percentages calculated from the level set by the initial test are misleading and should not be used as the basis for judging the student. Pre-study scores may best be used as the basis for a long-term study of how well such scores will predict post-study scores. Provisions of special classes might well be made for promising students, and possibly those whose predictive index is very low should be prevented from wasting their time and the teacher's energies. Whatever levels are set for inclusion or exclusion, they should be defined in terms of course and college objectives.

The Committee on Tests of the American Association of Physics Teachers

H. W. Farwell,
Harvey B. Lemon,
C. J. Lapp, *Chairman*,
Frederic Palmer, Jr.,
John T. Tate,
A. G. Worthing.